

빅데이터 시스템 기말 프로젝트

주제 : 우리나라 사람들의 독서 실태

이름 : 구명회
학과 : 컴퓨터공학과
학번 : 20154215
과목 : 빅데이터 시스템

INDEX_

- 독서의 필요성
- 구현한 기능 LIST
- 실험 결과
- 평가/고찰/기대효과
- 참고문헌과 소감

독서의 필요성?

독서는 뇌가 새로운 능력을 학습, 지능을 어떻게 확대하는지 명확히 보여준다. 던지기가 사냥을 위한 고도의 신체 기술이었듯 읽기는 뇌 속 소프트웨어의 성능을 향상시키는 신생 기술이다



- 언어의 기억 폭이 커진다.
- 사망 확률이 감소한다.
- 뇌 발달에 도움이 된다.
- 치매 예방에 월등하다
- 스트레스 해소에 도움이 된다.
- 독서를 하면 행복해진다.
- 우울증을 예방할 수 있다.
- 다이어트에 도움이 된다.

프로그램을 구현한 순서

- Project-1. 모듈 설치
- Project-2. 2018년 여가활동 응답 조사하기
- Project-3. 성인 남녀가 책을 읽는 비율 조사하기
- Project-4. 인구 1인당 평균 독서 권수 조사하기
- Project-5. 인구 1인당 평균 독서 권수에 대한 결론
- Project-6. 독서율이 줄어드는 이유 알아보기
- Project-7. 도서관 이용률 조사하기
- Project-8. 도서관의 위치 알아보기

Project-1. 모듈 설치

빅데이터 시스템 - 기말 프로젝트

주제 : 우리나라 사람들의 독서 실태

독서라 함은 일반도서를 종이책 또는 전자책으로 읽는 것을 말합니다. (일반도서란 웹소설이나 장르소설까지 포함합니다.)

통계청의 2017년 데이터 통계에 따르면 우리들은 책을 읽는 빈도가 한달에 한번꼴조차 되지 않습니다.

저는 이것의 원인을 규명하고 문제를 해결할 수 있는 방안을 생각해보려고 합니다.

Project-1. 모듈설치

1) pandas 모듈을 import 합니다.

통계청에서 다운로드 받은 데이터를 pandas 모듈을 사용해서 데이터 프레임 형식으로 저장할 것입니다.

2) matplotlib 모듈과 platform 모듈을 import 합니다.

데이터 프레임 형식으로 저장한 데이터를 matplotlib 모듈에서 지원하는 그래프를 사용해서 분석하려고 합니다.

그래프를 정상적으로 출력해주기 위해 %matplotlib inline 옵션을 설정하고, 한글 폰트를 적용해주었습니다.

3) 지도에 도서관의 위치를 표시하기 위해 folium 모듈과 googlemaps 모듈을 import 해주었습니다.

```
In [1]: import pandas as pd

import platform
import matplotlib.pyplot as plt
%matplotlib inline

path = "c:/Windows/Fonts/malgun.ttf"
from matplotlib import font_manager, rc
if platform.system() == 'Darwin':
    rc('font', family='AppleGothic')
elif platform.system() == 'Windows':
    font_name = font_manager.FontProperties(fname=path).get_name()
    rc('font', family=font_name)
else:
    print('Unknown system... sorry~~~~')
plt.rcParams['axes.unicode_minus'] = False

import folium
import googlemaps
```

Project-2 2018년 여가활동 복수응답 조사하기 (단위 : %)

Project-2. 2018년 여가활동 복수응답 조사하기 (단위 : %)

2018년 여가활동의 비율을 조사하려고 합니다.

pandas 모듈을 사용해서 play.xlsx를 데이터 프레임으로 저장했습니다.

다운로드 받은 데이터는 성별, 연령, 학력, 동거가구원수, 혼인상태별 등을 대상으로 조사한 2018년의 여가활동에 대한 응답을 저장하고 있습니다.

저는 이 중에서 간단하게 연령과 가구소득에 대해서만 그래프로 비교해보려고 합니다.

1) pandas 모듈을 사용해서 play 데이터 파일을 읽어왔습니다.

2) fillna() 옵션을 사용해서 데이터 파일의 존재하지 않는 Nan값을 제거하고 rename() 옵션을 사용해서 데이터 파일의 열 이름을 수정했습니다.

[4] : 인터넷 검색, 채팅, ucc 제작, sns 등

[10] : 독서(만화책, 웹툰 포함)

[15] : 어학/기술/자격증취득/학원

[16] : 캠블/복권 구입

[18] : 이색/테마카페체험

[19] : 보드게임, 퍼즐, 큐브 맞추기

[22] : 미니홈피, 블로그 관리

3) 데이터 프레임을 다듬기 위해 일부의 행과 열을 제거했습니다.

```
In [38]: play = pd.read_excel('../data/play.xlsx', header = 1)
play.fillna(method='pad', inplace=True)

play = play.rename(columns =
                    { play.columns[0] : '구분',
                      play.columns[4] : '인터넷',
                      play.columns[10] : '독서',
                      play.columns[15] : '자격증/학원',
                      play.columns[18] : '테마 카페 체험',
                      play.columns[1] : '분포',
                      play.columns[8] : '노래방',
                      play.columns[11] : '요리/다도',
                      play.columns[16] : '복권',
                      play.columns[19] : '보드게임', play.columns[21] : '블로그'}) 

del play['사례수']
play = play.drop([3, 60, 61, 62],0)
play.head(20)
```

Out [38]:

	구분	분포	쇼핑/외식	인터넷	음주	미용	게임	노래방	등산	독서	...	낚시	자격증/학원	복권	바둑, 장기, 체스	테마 카페 체험	보드게임	인테리어	블로그	미니홈피, 블로그관리	수집활동
0	전체	소계	85.5	70.4	59.1	48.3	38.0	31.3	30.9	24.0	...	11.0	8.8	6.6	6.0	5.4	4.9	4.1	4.1	3.4	1.8
1	성별	남성	81.2	73.3	74.4	32.7	47.0	32.4	38.5	24.0	...	20.0	9.6	8.2	11.0	5.3	5.6	3.0	1.0	3.1	2.0
2	성별	여성	89.7	67.6	43.9	63.9	29.0	30.2	23.3	24.0	...	2.0	7.9	5.1	1.0	5.6	4.2	5.3	7.3	3.6	1.6
4	연령	20대	86.5	86.3	72.0	51.3	71.1	51.6	11.4	47.3	...	4.0	27.6	5.0	1.8	15.6	12.2	4.9	4.2	8.1	2.7
5	연령	30대	91.1	85.8	68.5	48.5	53.6	34.9	21.8	28.1	...	9.8	6.9	5.4	2.4	6.2	5.6	7.0	5.7	6.7	2.5
6	연령	40대	88.8	82.0	67.1	50.5	33.8	29.9	38.3	20.3	...	14.4	3.9	6.1	4.8	2.5	3.0	5.2	4.4	2.0	2.0
7	연령	50대	87.5	72.9	64.1	50.9	22.0	27.3	51.2	13.8	...	17.4	2.2	7.5	8.1	1.5	1.9	4.1	4.4	1.4	1.6
8	연령	60대	84.3	47.9	53.8	48.2	13.9	20.0	46.5	9.0	...	13.8	0.9	9.8	10.2	1.6	0.8	2.1	3.5	0.7	0.7
9	연령	70대이상	73.4	17.4	39.3	41.7	6.2	7.5	20.6	5.4	...	7.7	0.5	10.2	12.6	1.4	1.0	0.9	2.2	0.0	0.6
10	학력	초졸 이하	73.8	26.4	32.6	49.9	14.8	16.2	16.8	10.8	...	6.3	2.7	11.6	8.5	1.7	3.7	0.7	2.8	0.3	0.5
11	학력	중졸	77.3	50.2	30.1	40.9	34.5	28.0	21.1	25.3	...	6.8	12.4	6.0	5.8	5.6	5.3	1.4	2.4	1.9	1.4
12	학력	고졸	87.0	70.6	63.0	51.2	36.1	33.0	33.5	22.8	...	11.4	8.5	6.1	5.5	5.4	4.2	3.4	3.7	2.6	1.3
13	학력	대졸이상	88.6	85.3	69.4	47.2	45.7	33.0	33.6	27.4	...	12.8	8.9	6.6	6.1	6.1	5.9	6.6	5.5	5.4	2.8
14	동거가구원수	1인	83.7	53.2	57.7	46.7	31.1	23.6	25.8	17.6	...	8.8	5.1	8.4	7.5	4.9	3.9	2.8	4.1	1.9	1.2
15	동거가구원수	2인	82.9	48.5	51.6	47.0	17.7	18.6	34.4	11.6	...	11.6	2.0	10.0	8.9	2.1	1.4	3.1	3.0	1.2	1.3
16	동거가구원수	3인이상	86.3	77.9	61.0	48.9	43.6	35.3	30.8	27.8	...	11.1	10.8	5.6	5.1	6.2	5.9	4.6	4.4	4.0	2.0
17	혼인상태별	미혼	84.1	86.8	58.7	47.6	71.8	50.2	13.6	47.7	...	5.5	25.2	4.6	2.4	13.6	12.2	4.1	4.2	7.8	2.7
18	혼인상태별	기혼	87.1	68.5	61.5	49.1	27.2	25.4	38.8	16.2	...	13.7	2.8	7.4	7.2	2.4	2.2	4.6	4.2	1.9	1.6
19	혼인상태별	사별/이혼/기타	78.2	33.2	42.7	45.6	11.1	14.7	27.3	7.1	...	7.8	0.9	7.4	8.5	2.3	1.9	1.1	3.4	0.3	0.7
20	가구주여부	가구주	82.8	64.9	72.2	37.1	32.4	24.9	42.0	15.7	...	20.6	3.7	9.1	11.3	3.1	2.8	3.1	1.6	2.1	1.8

20 rows × 23 columns

Project-2 2018년 여가활동 복수응답 조사하기 (단위 : %)

4) 읽어온 데이터파일에서 연령만 선택한 다음 데이터 프레임을 예쁘게 다듬기 위해 구분을 제거하고, 분포를 인덱스로 설정했습니다.

5) 수정한 데이터 프레임에서 몇 개의 속성만 선택해서 그래프로 그렸습니다.

6) 읽어온 데이터파일에서 가구소득만 선택한 다음 데이터 프레임을 예쁘게 다듬기 위해 구분을 제거하고, 분포를 인덱스로 설정했습니다.

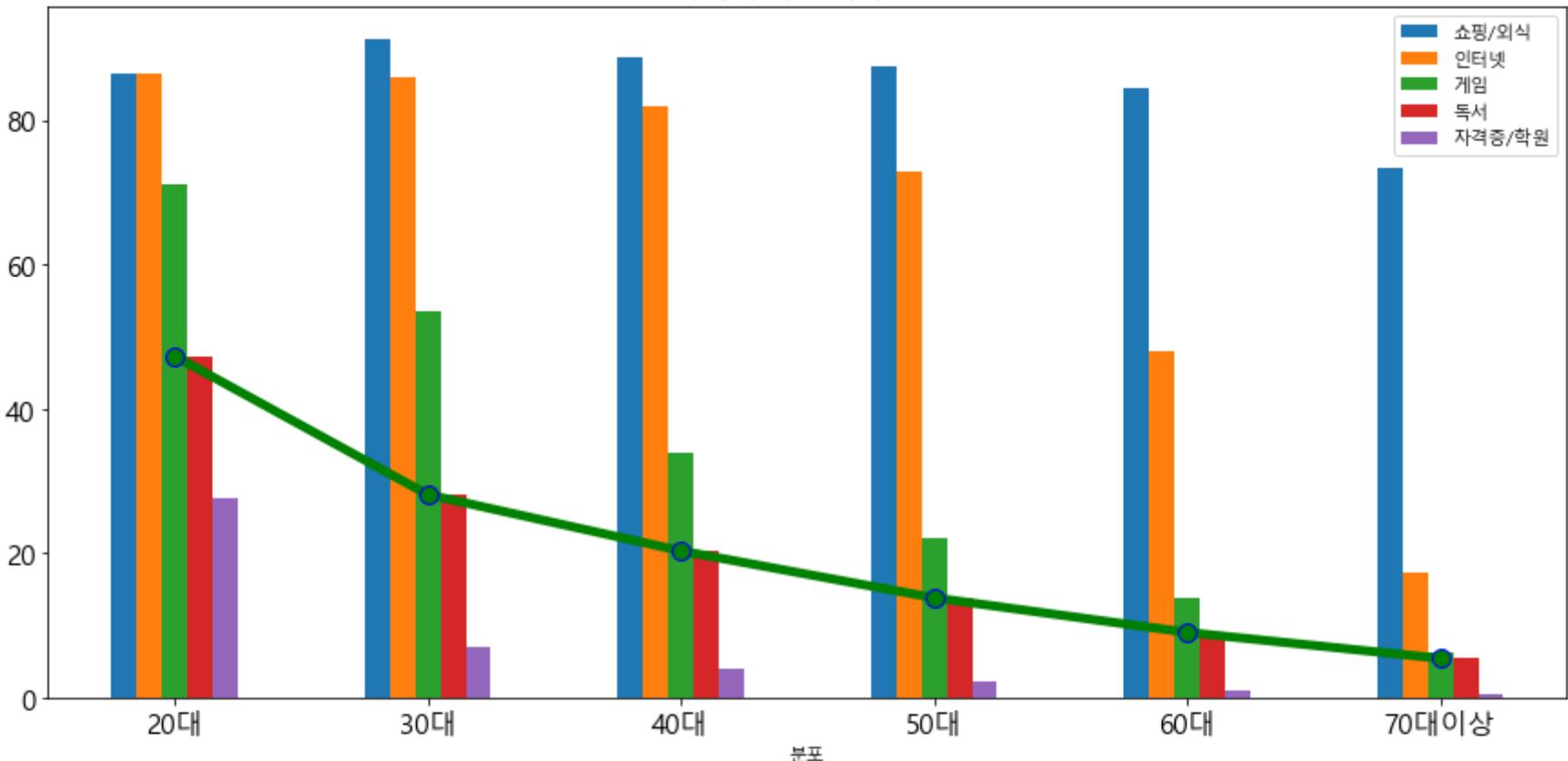
7) 수정한 데이터 프레임에서 몇 개의 속성만 선택해서 그래프로 그렸습니다.

결과 : 독서가 의외로 준수한 비율을 차지하고 있는 것처럼 보입니다.

하지만 이 독서의 비율은 앞에서 언급했던 순수한 독서의 비율이 아니라 만화나 웹툰 등을 포함한 비율입니다.

그것을 고려해보면 이는 절대로 높은 수치가 아닐 것입니다.

나이에 따른 여가활동

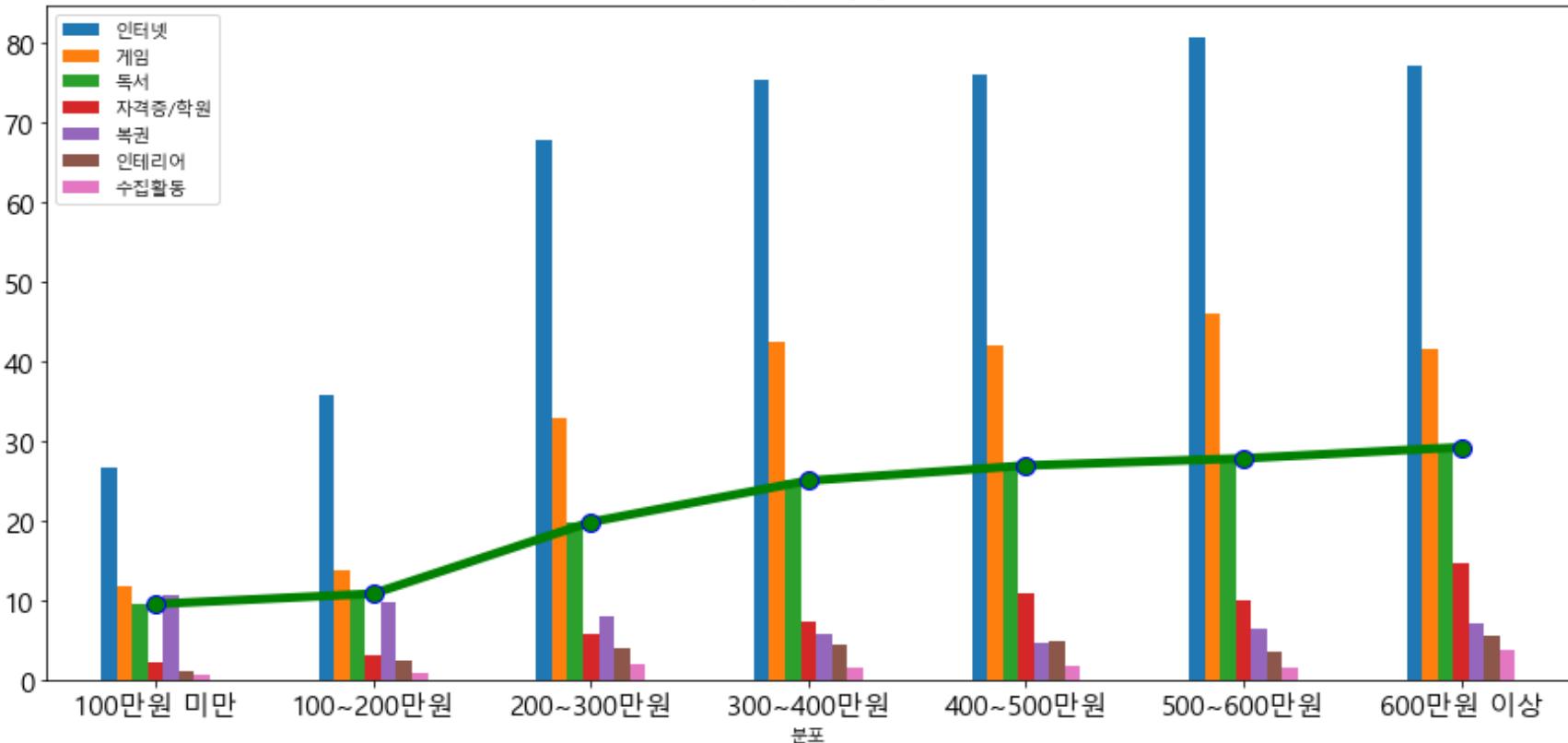


Out [31] :

쇼핑/외식 인터넷 게임 독서 자격증/학원

분포					
20대	86.5	86.3	71.1	47.3	27.6
30대	91.1	85.8	53.6	28.1	6.9
40대	88.8	82.0	33.8	20.3	3.9
50대	87.5	72.9	22.0	13.8	2.2
60대	84.3	47.9	13.9	9.0	0.9
70대이상	73.4	17.4	6.2	5.4	0.5

소득에 따른 여가활동



Out [35]:

	인터넷	게임	독서	자격증/학원	복권	인테리어	수집활동
분포							
100만원 미만	26.6	11.8	9.5	2.2	10.6	1.0	0.6
100~200만원	35.8	13.8	10.8	3.0	9.8	2.3	0.8
200~300만원	67.9	32.9	19.8	5.8	8.0	4.0	1.9
300~400만원	75.4	42.4	25.0	7.2	5.7	4.3	1.5
400~500만원	76.1	42.0	26.9	10.8	4.7	4.9	1.8
500~600만원	80.6	45.9	27.8	9.9	6.4	3.5	1.5
600만원 이상	77.2	41.6	29.2	14.7	7.1	5.5	3.7

Project-3.1. 성인 남녀가 종이책을 읽는 비율(단위 %)

Project-3.1. 성인 남녀가 종이책을 읽는 비율

2017년 성인 남녀가 종이책을 읽을 비율을 확인하려고 합니다.

pandas 모듈을 사용해서 paper_person.xlsx를 데이터 프레임으로 저장했습니다.

다운로드 받은 데이터는 2017년의 성별, 연령별, 학력별, 가구소득별로 종이책을 읽었는지 안 읽었는지 조사한 것을 저장하고 있습니다.

이번에는 성인 남녀의 데이터만 그래프로 비교해보려고 합니다.

1) pandas 모듈을 사용해서 paper_person 데이터 파일을 읽어왔습니다.

2) fillna() 옵션을 사용해서 데이터 파일의 존재하지 않는 Nan값을 제거하고 rename() 옵션을 사용해서 데이터 파일의 열 이름을 수정했습니다.

3) 읽어온 데이터파일에서 성별만 선택한 다음 데이터 프레임을 예쁘게 다듬기 위해 '구분'과 '사례수'를 제거하고, '분포'를 인덱스로 설정했습니다.

4) 수정한 데이터 프레임을 그래프로 그렸습니다.

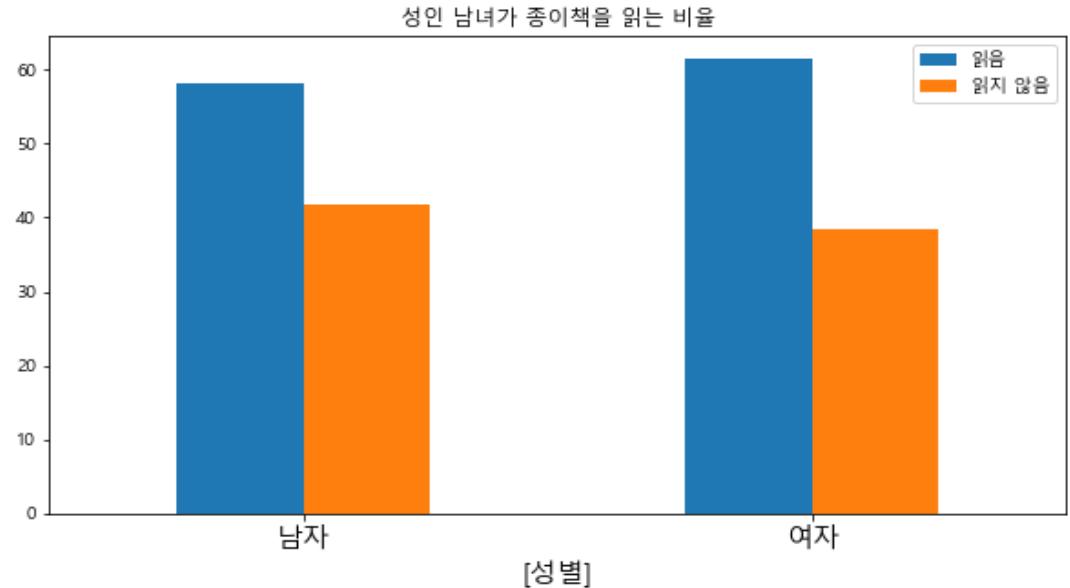
결과 : 남/녀 모두 읽음이 읽지 않음보다 20%정도 높은걸 알 수 있습니다.

```
In [5]: paper_person = pd.read_excel('../data/paper_person.xlsx', header = 1)
paper_person.fillna(method='pad', inplace=True)

paper_person = paper_person.rename(columns={paper_person.columns[0] : '구분', paper_person.columns[1] : '분포'})

paper_person = paper_person[paper_person['구분'] == '성별']
del paper_person['구분']
del paper_person['사례수']
paper_person.set_index('분포', inplace=True)

paper_person.plot(title = '성인 남녀가 종이책을 읽는 비율', kind='bar', figsize=(10,5))
plt.xticks(fontsize=15, rotation = 0)
plt.xlabel("[성별]", fontsize=15, rotation=0)
plt.show()
paper_person
```



Out [5] :

읽음 읽지 않음

분포	읽음	읽지 않음
남자	58.2	41.8
여자	61.5	38.5

Project-3.2. 성인 남녀가 종이책 또는 전자책을 읽는 비율(단위 %)

Project-3.2. 성인 남녀가 종이책 또는 전자책을 읽는 비율

2017년 성인 남녀가 종이책 또는 전자책을 읽을 비율을 확인하고 합니다.

pandas 모듈을 사용해서 paper_e_person.xlsx를 데이터 프레임으로 저장했습니다.

다운로드 받은 데이터는 2017년의 성별, 연령별, 학력별, 가구소득별로 책을 읽었는지 안 읽었는지 조사한 것을 저장하고 있습니다.

마찬가지로 성인 남녀의 데이터만 그래프로 비교하겠습니다.

1) pandas 모듈을 사용해서 paper_e_person 데이터 파일을 읽어왔습니다.

2) fillna() 옵션을 사용해서 데이터 파일의 존재하지 않는 Nan값을 제거하고 rename() 옵션을 사용해서 데이터 파일의 열 이름을 수정했습니다.

3) 읽어온 데이터파일에서 성별만 선택한 다음 데이터 프레임을 예쁘게 다듬기 위해 '구분'과 '사례수'를 제거하고, '분포'를 인덱스로 설정했습니다.

4) 수정한 데이터 프레임을 그래프로 그렸습니다.

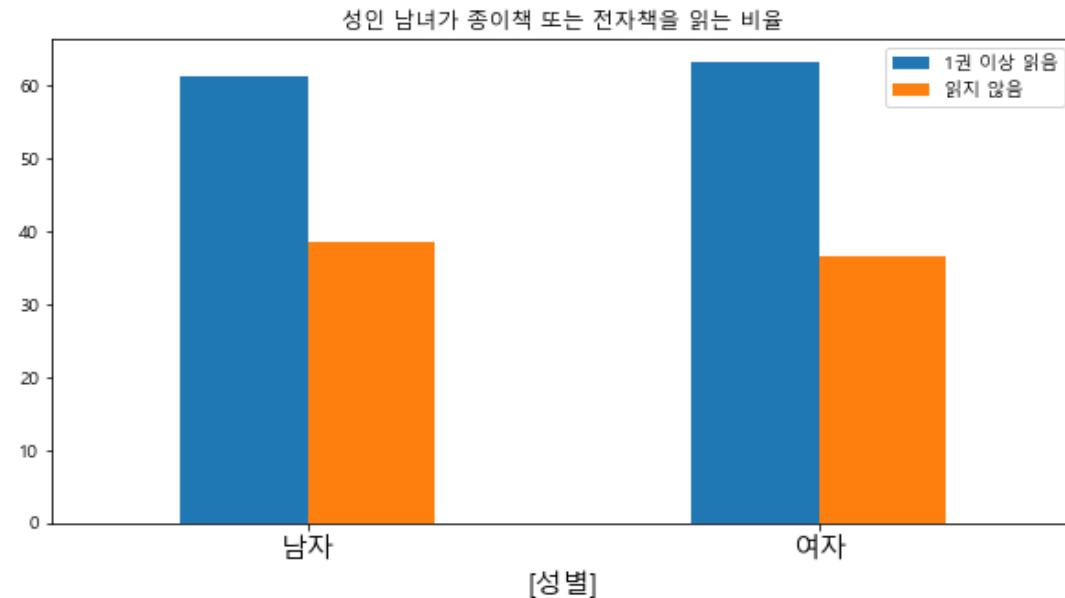
결과 : 남/녀 모두 읽음이 읽지 않음보다 25%정도 높은걸 알 수 있습니다.

```
In [10]: paper_e_person = pd.read_excel('../data/paper_e_person.xlsx', header = 1)
paper_e_person.fillna(method='pad', inplace=True)

paper_e_person = paper_e_person.rename(columns={paper_e_person.columns[0] : '구분', paper_e_person.columns[1] : '분포'})

paper_e_person = paper_e_person[paper_e_person['구분'] == '성별']
del paper_e_person['구분']
del paper_e_person['사례수']
paper_e_person.set_index('분포', inplace=True)

paper_e_person.plot(title = '성인 남녀가 종이책 또는 전자책을 읽는 비율', kind='bar', figsize=(10,5))
plt.xticks(fontsize=15, rotation = 0)
plt.xlabel("[성별]", fontsize=15, rotation=0)
plt.show()
paper_e_person
```



Out [10]:

1권 이상 읽음 읽지 않음

분포

분포	1권 이상 읽음	읽지 않음
남자	61.3	38.7
여자	63.3	36.7

Project-4. 인구 1인당 평균 독서 권수

Project-4. 인구 1인당 평균 독서 권수

이제 여가활동중에서 독서의 비율이 얼마정도나 차지하고 있는지 알았으니 이번에는 인구 1인당 평균 독서 권수를 확인하려고 합니다.

여기서 조사하는 독서는 만화나 웹툰 등을 제외한 순수한 일반도서를 읽는 것입니다.

pandas 모듈을 사용해서 reading_book.xlsx 파일을 데이터 프레임으로 저장했습니다.

다운로드 받은 데이터는 2009년부터 2017년까지 2년간격으로 조사한 인구 1인당 평균 독서 권수가 저장되어 있습니다.

합계, 잡지, 교양서적, 직업 관련 서적, 생활취미 정보 서적, 기타에 대한 column이 저장되어 있는데

하나하나를 비교하는 것도 재미있겠지만 프로젝트 진행에 관련이 없다고 생각해서 저는 이중에서 합계만 비교하도록 하겠습니다.

- 1) pandas 모듈을 사용해서 reading_book.xlsx 데이터 파일을 읽어왔습니다.
- 2) fillna() 옵션을 사용해서 데이터 파일의 존재하지 않는 Nan값을 제거했습니다.
- 3) 데이터 프레임을 예쁘게 다듬기 위해 필요없는 행들을 제거하고, reset_index() 옵션으로 인덱스를 초기화해서 재정렬시켰습니다.
- 4) rename() 옵션을 사용해서 데이터 파일의 열 이름을 수정했습니다.

```
In [121]: book = pd.read_excel('../data/reading_book.xlsx', header = 1, usecols = 'B, C, E, S, AG, AU, BI')
book.fillna(method='pad', inplace=True)

book = book.drop([1,2,43,44,45,46],0)
book.reset_index(drop=True, inplace=True)

book = book.rename(columns={
    book.columns[0] : '구분', book.columns[1] : '분포',
    book.columns[2] : '인구 1인당 평균 독서 권수(2009)',
    book.columns[3] : '인구 1인당 평균 독서 권수(2011)',
    book.columns[4] : '인구 1인당 평균 독서 권수(2013)',
    book.columns[5] : '인구 1인당 평균 독서 권수(2015)',
    book.columns[6] : '인구 1인당 평균 독서 권수(2017)'})

target_col = [ '인구 1인당 평균 독서 권수(2009)', '인구 1인당 평균 독서 권수(2011)',
    '인구 1인당 평균 독서 권수(2013)', '인구 1인당 평균 독서 권수(2015)', '인구 1인당 평균 독서 권수(2017)' ]

book.head(20)
```

Out [121]:

	구분	분포	인구 1인당 평균 독서 권수(2009)	인구 1인당 평균 독서 권수(2011)	인구 1인당 평균 독서 권수(2013)	인구 1인당 평균 독서 권수(2015)	인구 1인당 평균 독서 권수(2017)
0	전체	계	10.8	12.8	11.2	9.3	9.5
1	성별	남자	11.9	14.4	12.7	10.4	10.5
2	성별	여자	9.8	11.3	9.6	8.2	8.5
3	연령	20~29세	17.9	18.8	17.2	14.0	14.5
4	연령	30~39세	14.5	16.6	14.5	13.1	12.4
5	연령	40~49세	10.4	13.0	11.7	9.6	11.7
6	연령	50~59세	5.8	7.7	6.2	5.9	6.0
7	연령	60세 이상	4.4	5.3	4.9	5.1	5.3
8	교육정도	초졸이하	2.3	7.2	6.5	4.5	4.9
9	교육정도	중·졸	7.7	9.3	7.5	5.5	5.9
10	교육정도	고·졸	9.9	11.1	9.7	8.1	8.0
11	교육정도	대졸이상	17.7	19.3	16.5	14.0	13.9
12	혼인상태	미·혼	18.0	20.2	18.1	14.5	14.5
13	혼인상태	배우자 있음	9.1	10.8	9.1	7.7	8.3
14	혼인상태	사·별	1.7	2.2	1.8	1.8	1.7
15	혼인상태	이·혼	6.0	6.9	5.7	5.5	5.1
16	경제활동상태	취·업	11.2	13.1	11.1	9.6	9.9
17	경제활동상태	실업 및 비경제활동	10.3	12.5	11.3	8.8	9.0
18	산업	농림어업	2.0	2.7	2.8	2.3	2.1
19	산업	광공업	9.8	12.6	9.3	8.3	8.8

Project-4.1 인구 1인당 평균 독서 권수(연령)

Project-4.1 인구 1인당 평균 독서 권수(연령)

전체 데이터에서 연령 데이터만 선택해서 비교해보겠습니다.

1) '구분' 열의 '연령'에 해당하는 값들을 age에 저장하고, 인덱스를 '분포'로 재설정했습니다.

2) age 데이터에 대한 그래프를 그리고, 추가적으로 데이터 프레임을 출력했습니다.

결과 : 독서 평균이 40 ~ 49세 부터는 조금 올라갔거나 비슷하지만 20 ~ 39세에서는 전체적으로 내려갔음을 알 수 있습니다.

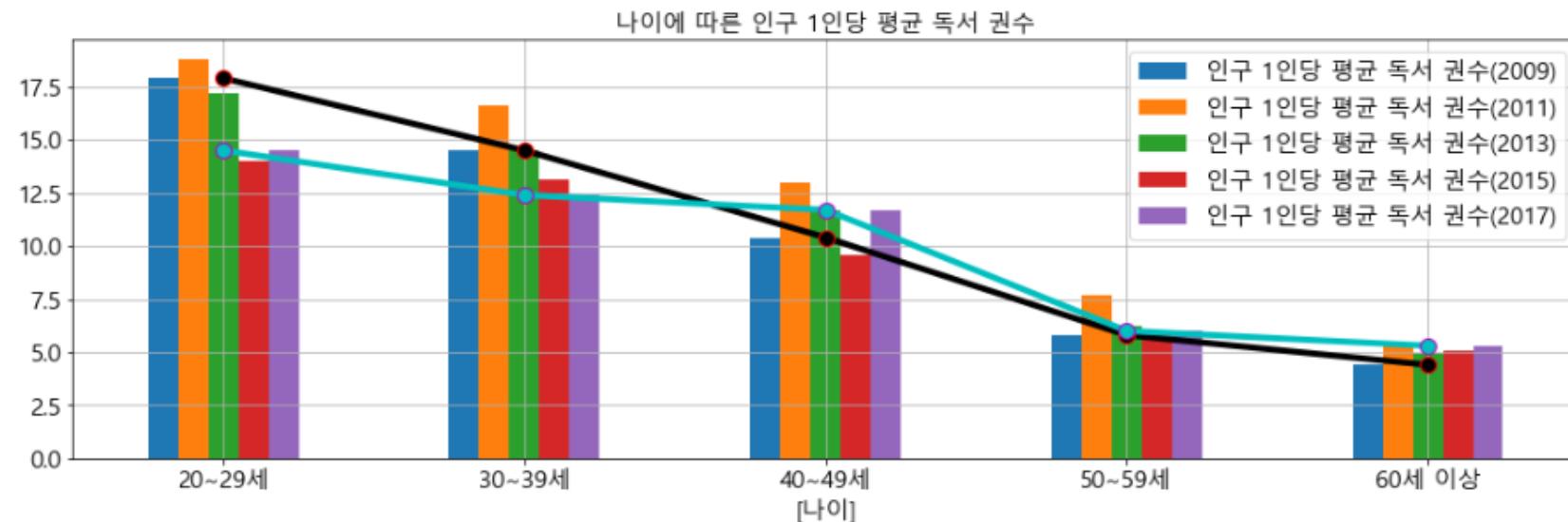
제 생각은 40 ~ 49세부터 그래프가 증가하는 이유는 애초에 표본이 20 ~ 39세보다 적기도 하지만, 보통 40세정도면 안정적인 생활을 할 것이고,

직업상의 이유나 취미를 찾는등 어떠한 경유로 책을 읽게되는 분들이 많아졌을것이라고 생각합니다.

구체적인 결과를 마지막에서 비교해보겠습니다.

```
In [8]: age = book[book['구분'] == '연령']
del age['구분']
age.set_index('분포', inplace=True)

age[target_col].plot(kind='bar', grid=True, figsize=(17,5))
plt.plot(age['인구 1인당 평균 독서 권수(2009)'], linewidth = 4, color='k', marker = 'o', ms = 10, mec = 'r')
plt.plot(age['인구 1인당 평균 독서 권수(2017)'], linewidth = 4, color='c', marker = 'o', ms = 10, mec = 'm')
plt.title("나이에 따른 인구 1인당 평균 독서 권수", fontsize=15)
plt.legend(fontsize = 15, loc='best')
plt.xlabel("[나이]", rotation=0, fontsize=15)
plt.xticks(fontsize = 15, rotation=0)
plt.yticks(fontsize = 15)
plt.show()
age
```



Out [8] :

분포	인구 1인당 평균 독서 권수(2009)	인구 1인당 평균 독서 권수(2011)	인구 1인당 평균 독서 권수(2013)	인구 1인당 평균 독서 권수(2015)	인구 1인당 평균 독서 권수(2017)
20~29세	17.9	18.8	17.2	14.0	14.5
30~39세	14.5	16.6	14.5	13.1	12.4
40~49세	10.4	13.0	11.7	9.6	11.7
50~59세	5.8	7.7	6.2	5.9	6.0
60세 이상	4.4	5.3	4.9	5.1	5.3

Project-4.2 인구 1인당 평균 독서 권수(연령*성별)

Project-4.2 인구 1인당 평균 독서 권수(연령*성별)

연령대별로 조사하는김에 연령대별로 성별데이터도 조사하려고 합니다. 아마 비슷한 분포를 보일것입니다.

전체 데이터에서 연령*성별 데이터를 선택해서 비교해보겠습니다.

1) '구분' 열의 '연령*성별'에 해당하는 값들을 age_sex에 저장하고, 인덱스를 '분포'로 재설정했습니다.

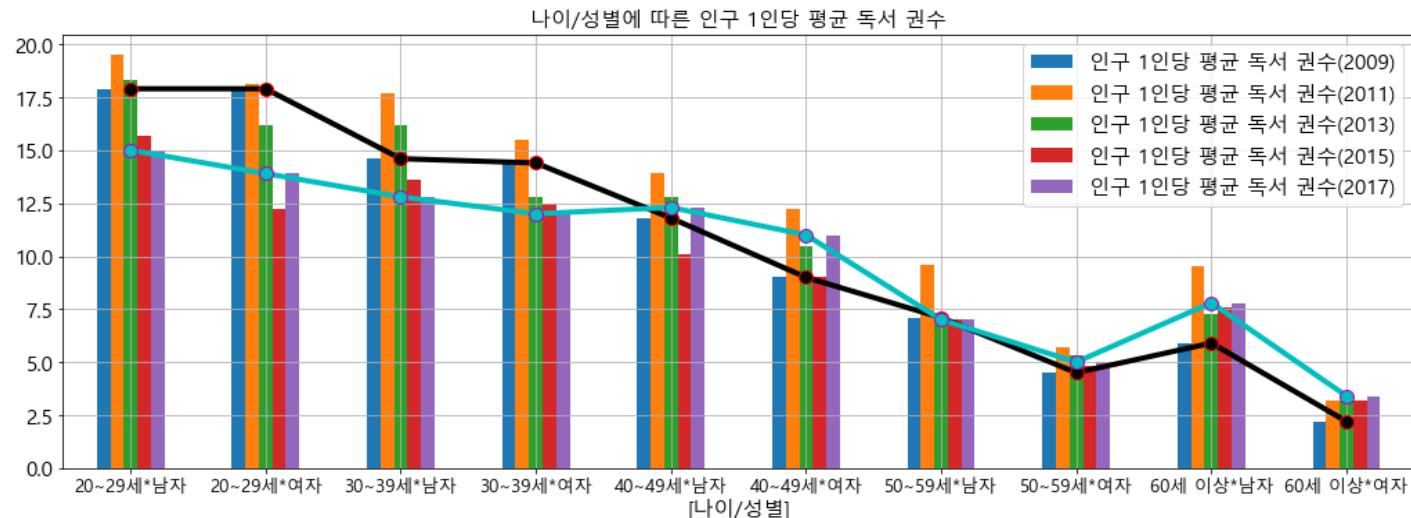
2) age_sex 데이터에 대한 그래프를 그리고, 추가적으로 데이터 프레임을 출력했습니다.

결과 : 예상대로 연령 데이터와 비슷한 결과가 나왔습니다. 대부분 남성이 여성보다 독서 권수가 많다는 것을 알 수 있습니다.

구체적인 결과를 마지막에서 비교해보겠습니다.

```
In [9]: age_sex = book[book['구분'] == '연령*성별']
del age_sex['구분']
age_sex.set_index('분포', inplace=True)

age_sex[target_col].plot(kind='bar', grid=True, figsize=(17.5,6))
plt.plot(age_sex['인구 1인당 평균 독서 권수(2009)'], linewidth = 4, color='k', marker = 'o', ms = 10, mec = 'r')
plt.plot(age_sex['인구 1인당 평균 독서 권수(2017)'], linewidth = 4, color='c', marker = 'o', ms = 10, mec = 'm')
plt.title("나이/성별에 따른 인구 1인당 평균 독서 권수", fontsize=15)
plt.legend(fontsize = 15, loc='best')
plt.xlabel('[나이/성별]', rotation=0, fontsize=15)
plt.xticks(fontsize = 13, rotation=0)
plt.yticks(fontsize = 15)
plt.show()
age_sex
```



Out [9]:

인구 1인당 평균 독서 권수(2009) 인구 1인당 평균 독서 권수(2011) 인구 1인당 평균 독서 권수(2013) 인구 1인당 평균 독서 권수(2015) 인구 1인당 평균 독서 권수(2017)

분포

분포	인구 1인당 평균 독서 권수(2009)	인구 1인당 평균 독서 권수(2011)	인구 1인당 평균 독서 권수(2013)	인구 1인당 평균 독서 권수(2015)	인구 1인당 평균 독서 권수(2017)
20~29세*남자	17.9	19.5	18.3	15.7	15.0
20~29세*여자	17.9	18.1	16.2	12.2	13.9
30~39세*남자	14.6	17.7	16.2	13.6	12.8
30~39세*여자	14.4	15.5	12.8	12.5	12.0
40~49세*남자	11.8	13.9	12.8	10.1	12.3
40~49세*여자	9.0	12.2	10.5	9.0	11.0
50~59세*남자	7.1	9.6	7.0	7.0	7.0
50~59세*여자	4.5	5.7	5.3	4.8	5.0
60세 이상*남자	5.9	9.5	7.3	7.6	7.8
60세 이상*여자	2.2	3.2	3.2	3.2	3.4

Project-4.3 인구 1인당 평균 독서 권수(교육정도)

Project-4.3 인구 1인당 평균 독서 권수(교육정도)

전체 데이터에서 교육정도에 대한 데이터만 선택해서 비교해보겠습니다.

1) '구분' 열의 '교육정도'에 해당하는 값들을 edu에 저장하고, 인덱스를 '분포'로 재설정했습니다.

2) edu 데이터에 대한 그래프를 그리고, 추가적으로 데이터 프레임을 출력했습니다.

결과 : 독서를 하는 정도가 고등 교육으로 올라갈수록 높아지는 것을 알 수 있습니다.

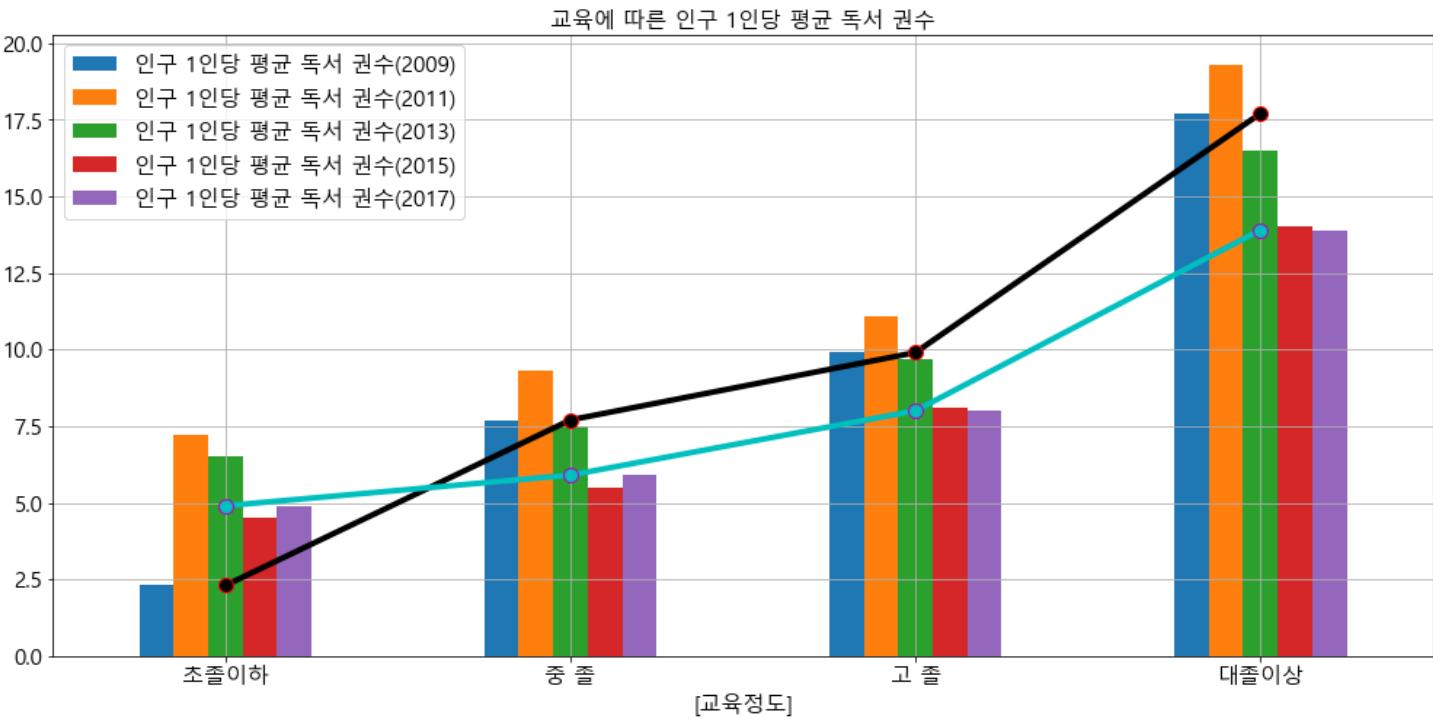
제 생각은 물론 고등 교육을 받지 않았다고 서적을 읽지 못한다는건 아니지만 상대적으로 고등 교육을 받을수록

교육과 관련된 교양 서적이나 전문 서적을 접하게 될 기회가 생기기 때문에 필연적인 결과라고 생각합니다.

구체적인 결과를 마지막에서 비교해보겠습니다.

```
In [10]: edu = book[book['구분'] == '교육정도']
del edu['구분']
edu.set_index('분포', inplace=True)

edu[target_col].plot(kind='bar', grid=True, figsize=(17,8))
plt.plot(edu['인구 1인당 평균 독서 권수(2009)'], linewidth = 4, color='k', marker = 'o', ms = 10, mec = 'r')
plt.plot(edu['인구 1인당 평균 독서 권수(2017)'], linewidth = 4, color='c', marker = 'o', ms = 10, mec = 'm')
plt.title("교육에 따른 인구 1인당 평균 독서 권수", fontsize=15)
plt.legend(fontsize = 15, loc='best')
plt.xlabel("[교육정도]", rotation=0, fontsize=15)
plt.xticks(fontsize = 15, rotation=0)
plt.yticks(fontsize = 15)
plt.show()
edu
```



Out [10]:

인구 1인당 평균 독서 권수(2009) 인구 1인당 평균 독서 권수(2011) 인구 1인당 평균 독서 권수(2013) 인구 1인당 평균 독서 권수(2015) 인구 1인당 평균 독서 권수(2017)

분포

초졸이하	2.3	7.2	6.5	4.5	4.9
중 졸	7.7	9.3	7.5	5.5	5.9
고 졸	9.9	11.1	9.7	8.1	8.0
대졸이상	17.7	19.3	16.5	14.0	13.9

Project-4.4 인구 1인당 평균 독서 권수(혼인상태)

Project-4.4 인구 1인당 평균 독서 권수(혼인상태)

전체 데이터에서 혼인상태에 대한 데이터만 선택해서 비교해보겠습니다.

1) '구분' 열의 '혼인상태'에 해당하는 값을 merry에 저장하고, 인덱스를 '분포'로 재설정했습니다.

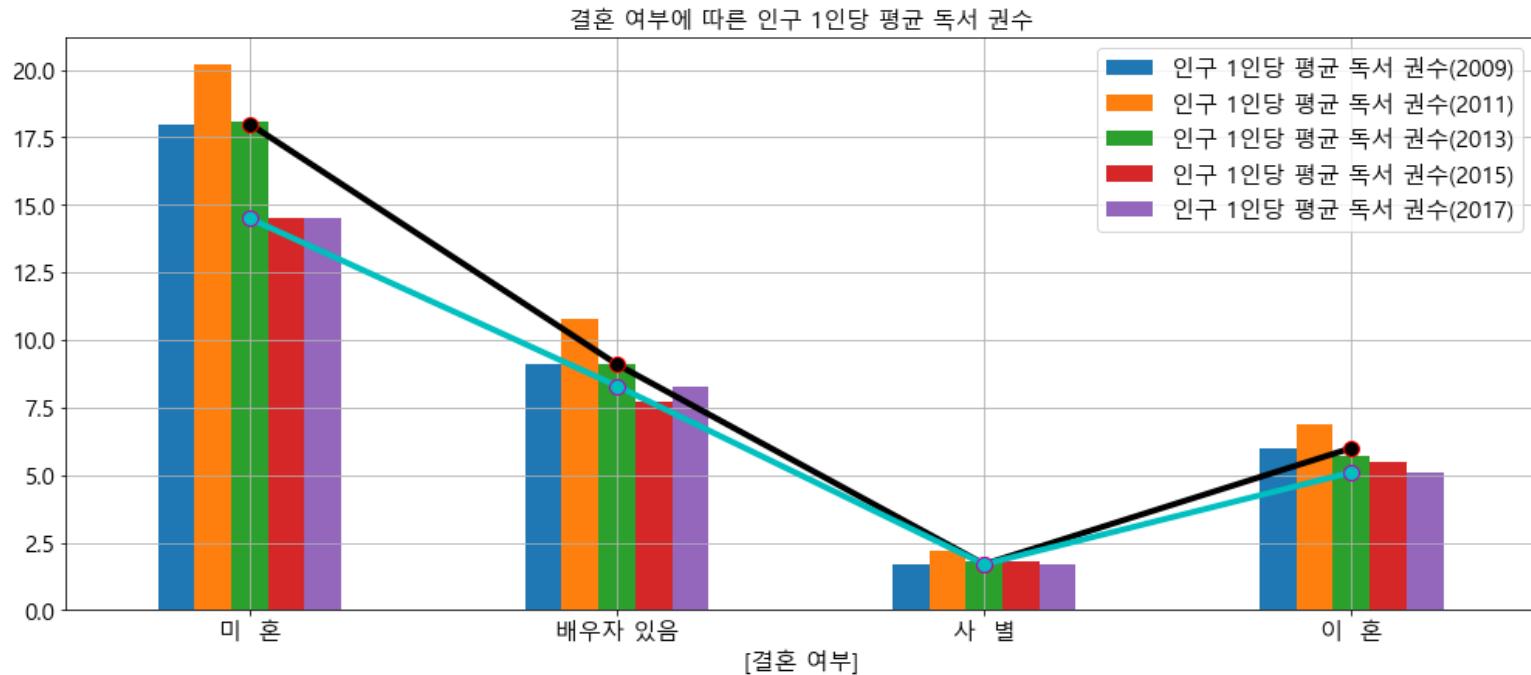
2) merry 데이터에 대한 그래프를 그리고, 추가적으로 데이터 프레임을 출력했습니다.

결과 : 결혼하고 나면 인생의 여유가 없어지는지, 결혼을 하고나면 독서율이 급감하는 것을 알 수 있습니다.

구체적인 결과를 마지막에서 비교해보겠습니다.

```
In [11]: merry = book[book['구분'] == '혼인상태']
del merry['구분']
merry.set_index('분포', inplace=True)

merry[target_col].plot(kind='bar', grid=True, figsize=(17,7))
plt.title("결혼 여부에 따른 인구 1인당 평균 독서 권수", fontsize=15)
plt.plot(merry['인구 1인당 평균 독서 권수(2009)'], linewidth = 4, color='k', marker = 'o', ms = 10, mec = 'r')
plt.plot(merry['인구 1인당 평균 독서 권수(2017)'], linewidth = 4, color='c', marker = 'o', ms = 10, mec = 'm')
plt.legend(fontsize = 15, loc='best')
plt.xticks(fontsize = 15, rotation=0)
plt.yticks(fontsize = 15)
plt.xlabel('[결혼 여부]', rotation=0, fontsize=15)
plt.show()
merry
```



Out [11]:

인구 1인당 평균 독서 권수(2009) 인구 1인당 평균 독서 권수(2011) 인구 1인당 평균 독서 권수(2013) 인구 1인당 평균 독서 권수(2015) 인구 1인당 평균 독서 권수(2017)

분포	인구 1인당 평균 독서 권수(2009)	인구 1인당 평균 독서 권수(2011)	인구 1인당 평균 독서 권수(2013)	인구 1인당 평균 독서 권수(2015)	인구 1인당 평균 독서 권수(2017)
미 혼	18.0	20.2	18.1	14.5	14.5
배우자 있음	9.1	10.8	9.1	7.7	8.3
사 별	1.7	2.2	1.8	1.8	1.7
이 혼	6.0	6.9	5.7	5.5	5.1

Project-4.5 인구 1인당 평균 독서 권수(경제활동상태)

Project-4.5 인구 1인당 평균 독서 권수(경제활동상태)

전체 데이터에서 경제활동상태에 대한 데이터만 선택해서 비교해보겠습니다.

1) '구분' 열의 '경제활동상태'에 해당하는 값들을 work에 저장하고, 인덱스를 '분포'로 재설정했습니다.

2) work 데이터에 대한 그래프를 그리고, 추가적으로 데이터 프레임을 출력했습니다.

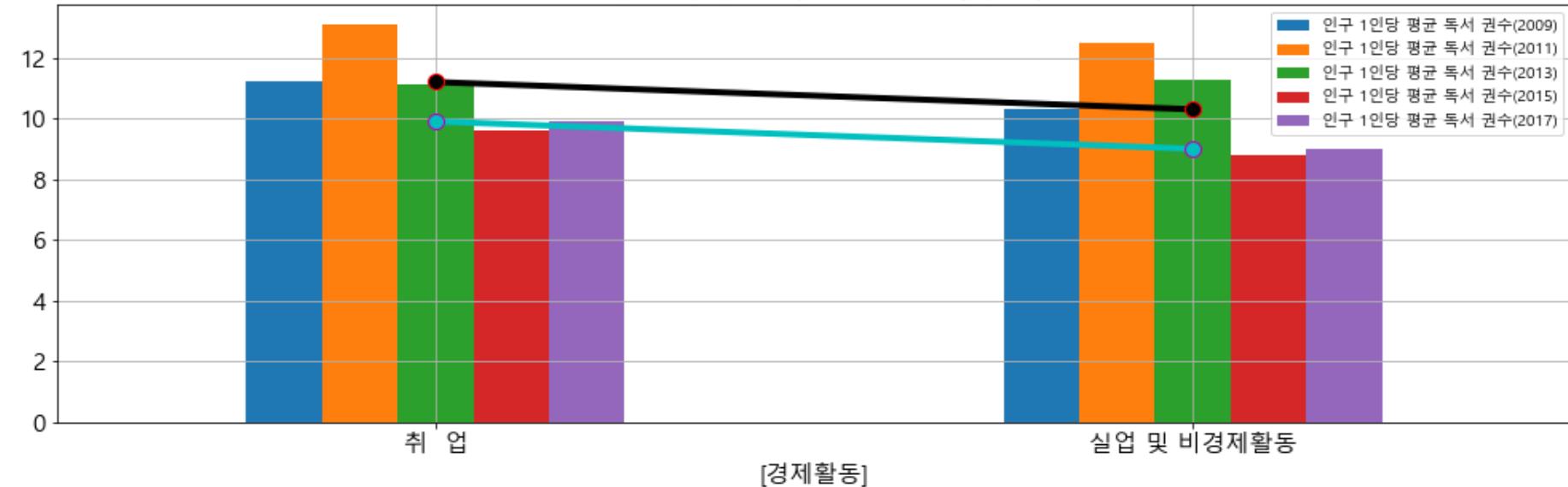
결과 : 경제활동상태에 따른 차이는 거의 없습니다.

구체적인 결과를 마지막에서 비교해보겠습니다.

```
In [12]: work = book[book['구분'] == '경제활동상태']
del work['구분']
work.set_index('분포', inplace=True)
```

```
work[target_col].plot(kind='bar', grid=True, figsize=(17,5))
plt.plot(work['인구 1인당 평균 독서 권수(2009)'], linewidth = 4, color='k', marker = 'o', ms = 10, mec = 'r')
plt.plot(work['인구 1인당 평균 독서 권수(2011)'], linewidth = 4, color='c', marker = 'o', ms = 10, mec = 'm')
plt.title("경제활동에 따른 인구 1인당 평균 독서 권수", fontsize=15)
plt.legend(fontsize = 10, loc='best')
plt.xlabel('[경제활동]', rotation=0, fontsize=15)
plt.xticks(fontsize = 15, rotation=0)
plt.yticks(fontsize = 15)
plt.show()
work
```

경제활동에 따른 인구 1인당 평균 독서 권수



Out [12]:

인구 1인당 평균 독서 권수(2009) 인구 1인당 평균 독서 권수(2011) 인구 1인당 평균 독서 권수(2013) 인구 1인당 평균 독서 권수(2015) 인구 1인당 평균 독서 권수(2017)

분포

분포	인구 1인당 평균 독서 권수(2009)	인구 1인당 평균 독서 권수(2011)	인구 1인당 평균 독서 권수(2013)	인구 1인당 평균 독서 권수(2015)	인구 1인당 평균 독서 권수(2017)
취업	11.2	13.1	11.1	9.6	9.9
실업 및 비경제활동	10.3	12.5	11.3	8.8	9.0

Project-4.6 인구 1인당 평균 독서 권수(가구소득)

Project-4.6 인구 1인당 평균 독서 권수(가구소득)

전체 데이터에서 가구소득에 대한 데이터만 선택해서 비교해보겠습니다.

1) '구분' 열의 '가구소득'에 해당하는 값을 money에 저장하고, 인덱스를 '분포'로 재설정했습니다.

2) money 데이터에 대한 그래프를 그리고, 추가적으로 데이터 프레임을 출력했습니다.

결과 : 금전적으로 여유가 있을수록 독서를 많이 하는것으로 나타났습니다.

구체적인 결과를 마지막에서 비교해보겠습니다.

In [13]:

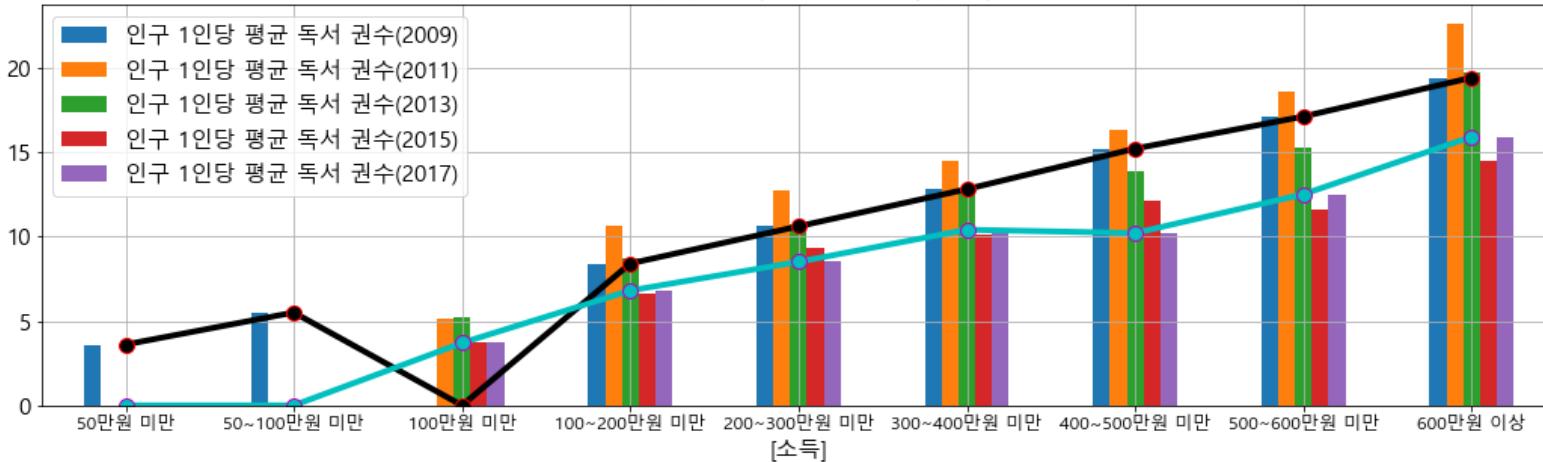
```

money = book[book['구분'] == '가구소득']
del money['구분']
money.set_index('분포', inplace=True)

money[target_col].plot(kind='bar', grid=True, figsize=(18,5))
plt.plot(money['인구 1인당 평균 독서 권수(2009)'], linewidth = 4, color='k', marker = 'o', ms = 10, mec = 'r')
plt.plot(money['인구 1인당 평균 독서 권수(2011)'], linewidth = 4, color='c', marker = 'o', ms = 10, mec = 'm')
plt.title("소득에 따른 인구 1인당 평균 독서 권수", fontsize=15)
plt.legend(fontsize = 15, loc='best')
plt.xlabel('[소득]', rotation=0, fontsize=15)
plt.xticks(fontsize = 12, rotation=0)
plt.yticks(fontsize = 15)
plt.show()
money

```

소득에 따른 인구 1인당 평균 독서 권수



Out [13]:

분포	인구 1인당 평균 독서 권수(2009)	인구 1인당 평균 독서 권수(2011)	인구 1인당 평균 독서 권수(2013)	인구 1인당 평균 독서 권수(2015)	인구 1인당 평균 독서 권수(2017)
50만원 미만	3.6	0.0	0.0	0.0	0.0
50~100만원 미만	5.5	0.0	0.0	0.0	0.0
100만원 미만	0.0	5.1	5.2	3.7	3.7
100~200만원 미만	8.4	10.6	8.7	6.6	6.8
200~300만원 미만	10.6	12.7	10.7	9.3	8.5
300~400만원 미만	12.8	14.5	12.7	10.1	10.4
400~500만원 미만	15.2	16.3	13.9	12.1	10.2
500~600만원 미만	17.1	18.6	15.3	11.6	12.5
600만원 이상	19.4	22.6	19.7	14.5	15.9

Project-5. 인구 1인당 평균 독서 권수에 대한 결론

Project-5. 인구 1인당 평균 독서 권수에 대한 결론

어떤 집단이던지 2011년에 독서량이 최고치를 달성하고, 그 다음부터 전체적으로 독서량이 줄어든 것을 확인할 수 있습니다.

마지막으로 2006년부터 2008년 데이터를 2016년과 2018년 여가활동 데이터와 비교해보겠습니다.

1) pandas 모듈을 사용해서 두 파일을 각각 데이터 프레임으로 저장하고 지금까지 했던 것처럼 데이터 프레임을 예쁘게 다듬었습니다.

다운로드 받은 데이터는 2006년부터 2008년까지의 여가활동에 대한 응답을 저장하고 있습니다.

2) 데이터를 비교할 몇 개의 열만 선택해서 비교해보겠습니다.

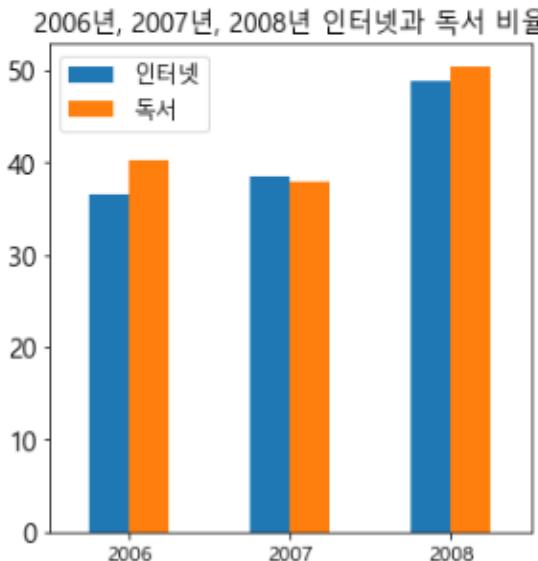
독서를 포함해서 비교적 높은 수치를 가지는 노래방, 게임, 인터넷, 쇼핑 등을 선택하겠습니다.

결론 : 인터넷의 비율과 독서의 비율이 반비례하고 있음을 알 수 있습니다.

여러가지 원인이 있겠지만 저는 독서의 비율이 줄어드는 가장 큰 원인이 스마트폰의 보급일 것이라고 생각합니다.

왜냐하면 2012년부터 스마트폰의 보급률이 급증하기 시작했고, 스마트폰 중독이라는 보도가 나타나기 시작한 시점이기 때문입니다.

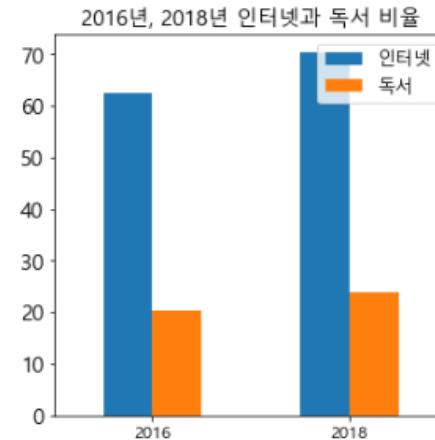
```
In [17]: find20062008[['인터넷', '독서']].plot(kind = 'bar', figsize=(5,5))
plt.title("2006년, 2007년, 2008년 인터넷과 독서 비율", fontsize=15)
plt.legend(fontsize = 13, loc='best')
plt.xlabel('', rotation=0, fontsize=15)
plt.xticks(fontsize = 12, rotation=0)
plt.yticks(fontsize = 15)
plt.show()
```



```
In [18]: find20162018 = pd.read_excel('../data/20162018.xlsx', header = 1)
find20162018.fillna(method='pad', inplace=True)
del find20162018['통계분류(1)']
del find20162018['사례수']

find20162018 = find20162018.rename(columns={ find20162018.columns[0] : '구분'})
find20162018.set_index('구분', inplace=True)

find20162018[['인터넷', '독서']].plot(kind = 'bar', figsize=(5,5))
plt.title("2016년, 2018년 인터넷과 독서 비율", fontsize=15)
plt.legend(fontsize = 13, loc='best')
plt.xlabel('', rotation=0, fontsize=15)
plt.xticks(fontsize = 12, rotation=0)
plt.yticks(fontsize = 15)
plt.show()
```



Project-5.1. 왜 독서율이 줄어드나?

Project-5.1. 왜 독서율이 줄어드나?

2013년, 2015년, 2017년의 독서량이 감소하는 이유를 조사한 통계가 있어 확인해보았습니다.

중간에 비어있는 데이터가 있어서 백분율을 다 채울 수 있도록 최대한 임의로 배분해서 작성했습니다.

임의로 작성한 부분은 []로 표시했고, 임의로 작성했기 때문에 그래프에서는 표시되지 않습니다.

결과 : 직장인은 스마트폰을 만질 시간조차 없는 것 같습니다. 일 때문에 시간이 없어서가 가장 큰 이유로 나왔고, 외외로 인터넷과 스마트폰의 영향이 적게 조사되었습니다.

하지만 "휴대전화 이용, 인터넷, 게임을 하느라" 항목이 두번째로 높게 집계된 걸 보면 영향이 없는 것은 아닌 것 같습니다.

추가적으로 비록 표본이 적지만, 주변 사람들에게 직접 물어본 결과 가장 많이 답했던 내용은 다음과 같습니다. (2019)

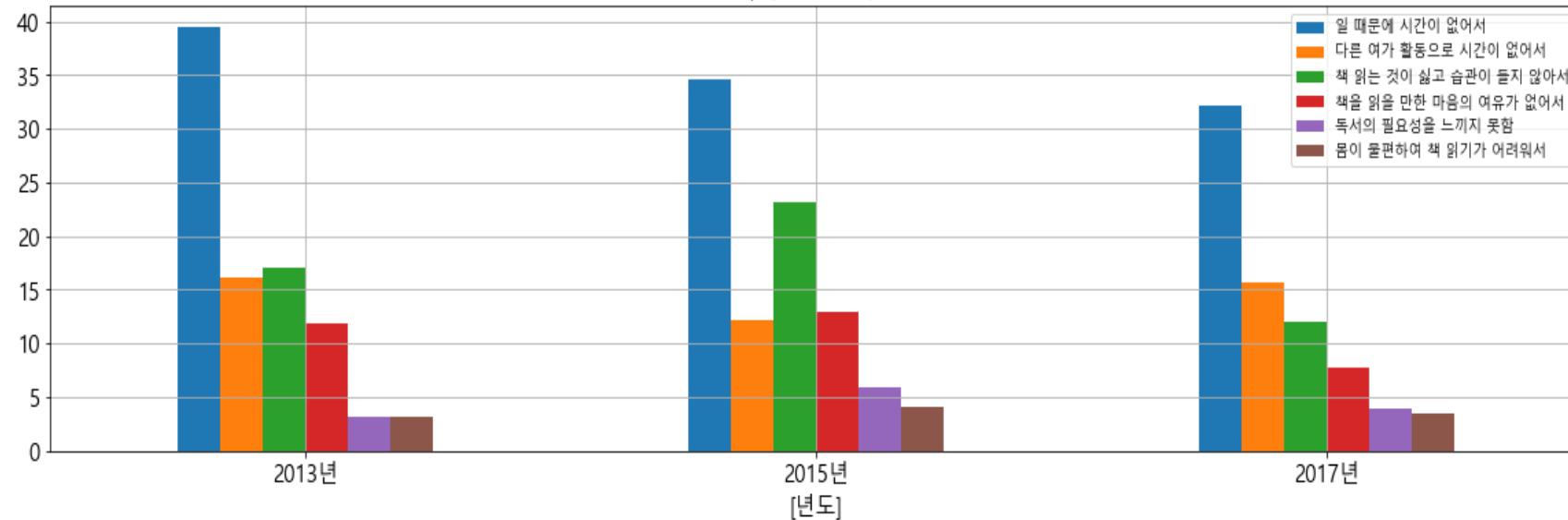
- 1) 책보다 유용한 정보를 얻을 수 있는 곳이 많기 때문입니다.
- 2) 책 살 돈과 시간을 인터넷을 검색하는데 투자하면 책에 자세히 나와있지 않은 다양한 부분까지 쉽게 구할 수 있습니다.
- 3) 현재 개인방송이 유행하고 있어 개인들이 전문지식을 뽐내며 녹화해서 투고하는 영상 데이터들도 많기 때문에 글자만 적혀있는 책을 보는 것보다 훨씬 이해가 잘 됩니다.
- 4) 특히 취업하고 회사를 다니는 경우 책을 찾아볼 시간이 부족한데, 책 대신 유튜브 등을 이용하면 시간도 절약되고 내용도 충분하기 때문에 책을 대체할 수 있다고 합니다.

```
In [21]: reading_reduce = pd.read_excel('../data/reading_reduce.xlsx', header = 0)
reading_reduce.fillna(method='pad', inplace=True)
reading_reduce.set_index('장애요인', inplace=True)
reading_reduce = reading_reduce.T

def reading_reduce['사례수']:
    reading_reduce.plot(kind='bar', grid=True, figsize=(20, 5))
    plt.title("독서율 감소 이유", fontsize=15)
    plt.legend(fontsize = 10, loc='best')
    plt.xlabel('[년도]', rotation=0, fontsize=15)
    plt.xticks(fontsize = 15, rotation=0)
    plt.yticks(fontsize = 15)
    plt.show()

reading_reduce
```

독서율 감소 이유



Out [21]:

장애요인	일 때문에 시간이 없어서	휴대전화 이용, 인터넷, 게임을 하느라	다른 여가 활동으로 시간이 없어서	책 읽는 것이 싫고 습관이 들지 않아서	책을 읽을 만한 마음의 여유가 없어서	독서의 필요성을 느끼지 못함	몸이 불편하여 책 읽기가 어려워서	어떤 책을 읽을지 모르겠다
2013년	39.5	[6.1]	16.1	17.1	11.8	3.2	3.1	3.1
2015년	34.6	[5.9]	12.2	23.2	12.9	5.9	4.1	[1.2]
2017년	32.2	19.6	15.7	12	7.7	4	3.5	[5.3]

Project-6. 도서관 이용률

Project-6. 도서관 이용률

한가지 흥미로운 사실이 있습니다.

관련 데이터가 없어서 확실하지 않지만, 예전에 얼핏 우리나라 공공 도서관의 수가 다른 나라와 비교했을 때 상위권에 속한다고 들은 적이 있습니다.

사실 독서를 도서관에서만 할 수 있는 것은 아니지만 도서관이야말로 책을 가장 많이 접할 수 있는 장소이기 때문에

저는 우리나라에서 공공 도서관을 이용하는 비율이 얼마나 되는지 찾아보기로 했습니다.

결과 : 100명중에 80명정도가 공공 도서관을 이용한 적이 없다는 정말 놀라운 결과를 발견할 수 있었습니다.

```
In [32]: library = pd.read_excel('../data/library.xlsx', header = 1)
library.fillna(method='pad', inplace=True)
library = library.rename(columns={library.columns[0] : '구분', library.columns[1] : '분포'})

del library['사례수']
del library['계']
library.set_index('구분', inplace=True)

library
```

Out [32]:

구분		분포	없음	1회 미만	1~5회	6~10회	11~20회	21회 이상	전체평균	이용자평균
전체	소계	77.8	8.7	10.8	2.0	0.6	0.0	0.5	2.5	
성별	남자	79.4	8.2	9.8	2.2	0.4	0.1	0.5	2.6	
성별	여자	76.3	9.3	11.8	1.9	0.7	0.0	0.6	2.4	
연령별	18~29세	57.8	15.1	18.9	6.1	2.2	0.0	1.4	3.3	
연령별	30대	71.3	9.3	16.8	2.1	0.5	0.1	0.6	2.2	
연령별	40대	76.5	10.2	11.7	1.3	0.3	0.2	0.5	2.0	
연령별	50대	86.4	6.8	6.2	0.4	0.1	0.0	0.2	1.6	
연령별	60대 이상	91.6	3.9	3.5	0.9	0.0	0.0	0.2	2.1	
학력별	중졸 이하	95.4	2.1	2.1	0.4	0.0	0.0	0.1	1.9	
학력별	고졸	87.6	5.8	5.6	0.7	0.2	0.0	0.2	2.0	
학력별	대재이상	64.5	13.2	17.7	3.5	1.0	0.1	0.9	2.6	
연간독서량	없음	96.1	2.6	0.9	0.4	0.1	0.0	0.1	1.8	
연간독서량	1~5권	69.3	14.5	13.3	2.5	0.4	0.0	0.6	1.9	
연간독서량	6~10권	69.7	11.0	16.4	2.3	0.6	0.1	0.7	2.4	
연간독서량	11~15권	53.9	13.5	25.3	4.4	2.9	0.0	1.4	3.1	
연간독서량	16~20권	73.3	9.9	13.9	2.0	1.0	0.0	0.6	2.3	
연간독서량	21권 이상	52.2	13.2	27.0	5.8	1.6	0.3	1.5	3.2	

Project-6-1. 도서관을 이용하지 않는 이유 조사하기

Project-6.1. 도서관을 이용하지 않는 이유 생각해보기

왜 이렇게 도서관을 이용하는 비율이 적을까 생각해보려다가 중요한것을 깨달았습니다.

정작 지금 조사하고 있는 저 역시 공공 도서관을 이용한 적이 없는 비율의 증가에 기여하고 있다는 사실입니다.

왜냐하면 도서관이 멀리있고, 조용한 분위기를 유지하면서 편하게 있을수도 없고, 차라리 그럴거면 집에서 하는게 편하고 집중도 잘되는데 굳이 멀리 있는 도서관에 갈 필요를 느끼지 못했기 때문입니다.

이제 2017년의 공공 도서관을 이용하지 않는 이유를 조사한 통계 데이터를 분석해보겠습니다.

결과 : 독서율이 줄어드는 이유와 비슷한 항목이 많이 보이며, 특히 '바빠서 이용할 시간이 없다'라는 항목이 1위를 차지했습니다.

이것을 보면 도서관의 이용률을 높이는 것이 독서율의 증가에 크게 도움이 되지 않을까 생각이 듭니다.

```
In [31]: nolibrary = pd.read_excel('../data/nolibrary.xlsx', header = 1)
```

```
nolibrary.fillna(method='pad', inplace=True)
```

```
nolibrary = nolibrary.rename(columns={nolibrary.columns[0] : '구분', nolibrary.columns[1] : '분포'})
```

```
del nolibrary['사례수']
```

```
del nolibrary['계']
```

```
nolibrary.set_index('구분', inplace=True)
```

```
nolibrary
```

```
Out[31]:
```

구분	분포	바빠서 이용할 시간 없다	이용할 필요성 느끼지 못한다	도서 대여점을 이용	이용 절차가 까다롭다	개관 시간이 짧다	항상 만원으로 좌석이 없다	독서상담 안내자 없음	읽고 싶은 책은 사서 읽는다	작장 도서실 이용한다	도서관 시설이 낡아서 쾌적하지 않다	도서관 이용 절차가 까다롭다
전체	소계	35.6		8.5	2.7	0.9	1.1	0.8	3.7	1.7	1.0	0.4
성별	남자	30.2		8.3	3.5	0.6	0.7	0.9	3.0	1.9	0.9	0.3
성별	여자	41.2		8.7	1.8	1.3	1.5	0.8	4.5	1.5	1.0	0.6
연령별	18~29세	20.5		9.3	3.3	2.2	1.4	1.0	5.9	7.0	1.6	0.5
연령별	30대	21.3		9.9	2.4	1.5	0.7	0.7	5.3	2.4	0.9	0.1
연령별	40대	24.9		7.1	2.7	0.9	0.7	0.8	4.3	0.6	1.1	0.5
연령별	50대	32.0		8.9	2.8	0.7	1.1	0.9	3.4	0.8	0.7	0.5
연령별	60대 이상	61.2		8.0	2.4	0.3	1.6	0.8	1.7	0.4	0.8	0.6
학력별	중졸 이하	67.4		6.7	2.4	0.2	1.4	0.0	0.9	0.4	0.0	0.5
학력별	고졸	35.7		8.2	2.3	1.0	1.3	1.1	3.3	0.6	1.3	0.5
학력별	대재이상	19.2		9.7	3.2	1.3	0.8	1.0	5.6	3.6	1.1	0.4
연간독서량	없음	42.3		5.9	2.5	0.7	1.4	0.8	1.9	0.8	0.5	0.5
연간독서량	1~5권	20.2		12.9	2.2	1.2	0.9	0.9	8.0	2.6	1.9	0.4
연간독서량	6~10권	34.0		10.8	3.7	1.6	0.3	0.1	3.0	2.0	0.8	0.6
연간독서량	11~15권	39.4		6.2	1.6	0.8	0.5	0.0	3.6	4.7	2.4	1.9
연간독서량	16~20권	39.1		6.9	2.3	1.0	1.7	3.4	5.0	0.7	0.5	0.0
연간독서량	21권 이상	33.8		11.5	3.6	0.6	1.1	0.1	4.4	4.4	1.4	0.0

Project-7. 도서관의 위치 알아보기

Project-7. 도서관의 위치 알아보기

마지막으로 우리나라 도서관이 얼마나 많은지 알아보고 근처에 있는 도서관을 지도상에 표기해보겠습니다.

저희집은 광산구에 위치하므로 우선 광주광역시 광산구의 모든 도서관을 찾아보고,

추가적으로 학교가 위치하는 광주광역시 등구의 모든 도서관도 찾아보도록 하겠습니다.

pandas 모듈을 사용해서 lotation.xls를 데이터 프레임으로 저장했습니다.

다운로드 받은 데이터는 전국의 도서관 이름과 주소, 위치정보 등이 저장되어 있습니다.

- 1) pandas 모듈을 사용해서 lotation.xls 데이터 파일을 읽어왔습니다.
- 2) fillna() 옵션을 사용해서 데이터 파일의 존재하지 않는 Nan값을 제거했습니다.
- 3) 읽어온 도서관 목록중에서 광주광역시 광산구와 등구에 위치하는 도서관 리스트들을 선택했습니다.
- 4) 데이터 프레임에 저장된 도서관들의 위도와 경도를 고유 google key를 이용해서 불러온 googlemap에 표기했습니다.

```
In [20]: location = pd.read_excel('../data/lotation.xls', header = 0)
location.fillna(method='pad', inplace=True)

location = location[location['시도명'] == '광주광역시']
location = location[location['시군구명'] == '광산구']
location.reset_index(drop=True, inplace=True)

location[['도서관명', '시도명', '시군구명', '도서관유형', '운영기관명', '도서관전화번호', '홈페이지주소', '위도', '경도']]
```

Out [20]:

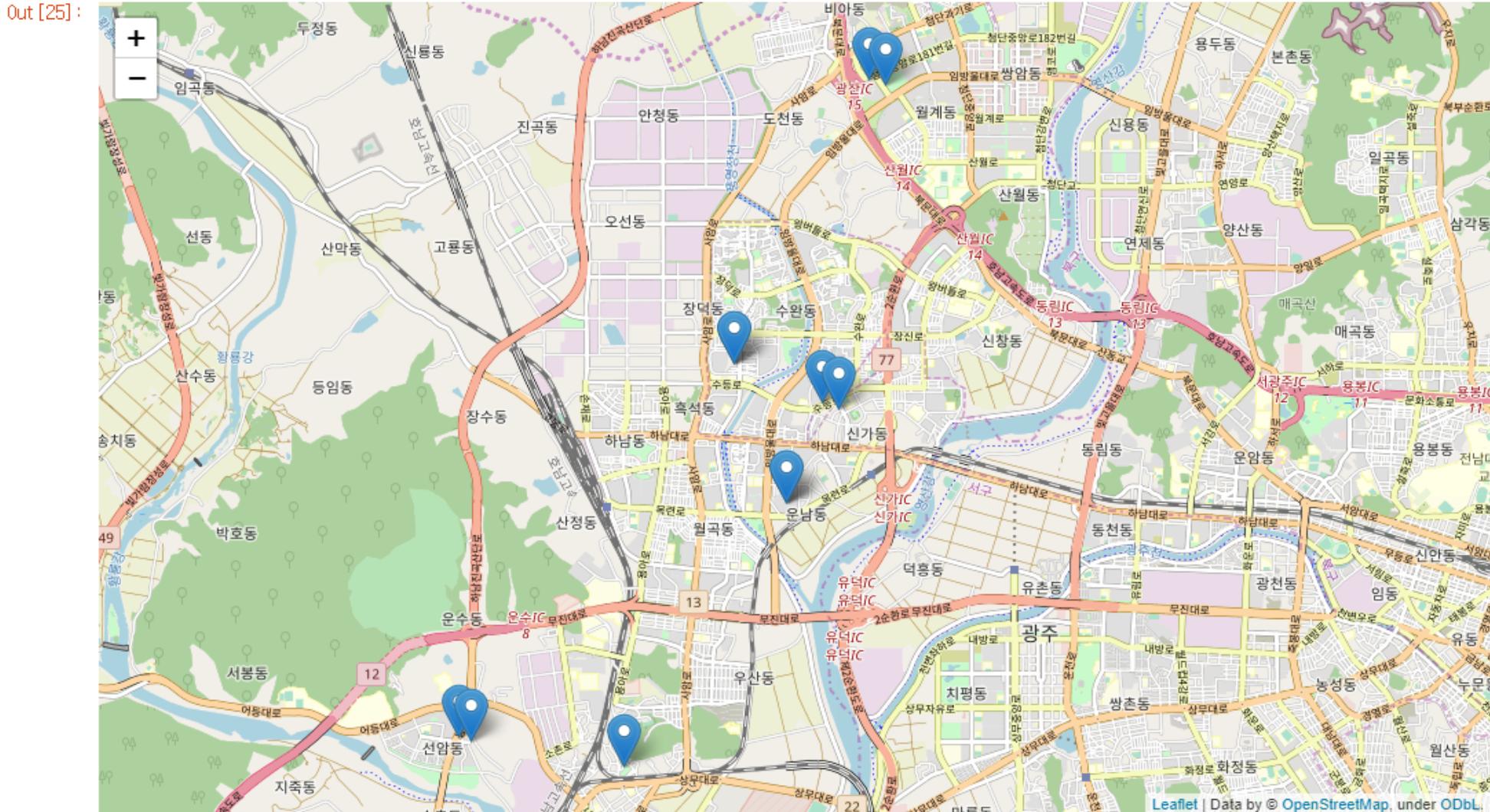
	도서관명	시도명	시군구명	도서관유형	운영기관명	도서관전화번호	홈페이지주소	위도	경도
0	신가도서관	광주광역시	광산구	공공도서관	광주광역시 광산구청	062-960-8514	http://lib.gwangsan.go.kr	35.182797	126.826386
1	운남어린이도서관	광주광역시	광산구	공공도서관	광주광역시 광산구청	062-960-8303	http://lib.gwangsan.go.kr	35.173025	126.819786
2	이야기꽃도서관	광주광역시	광산구	공공도서관	광주광역시 광산구청	062-960-6812	http://lib.gwangsan.go.kr	35.148414	126.778992
3	장덕도서관	광주광역시	광산구	공공도서관	광주광역시 광산구청	062-960-3989	http://lib.gwangsan.go.kr	35.187519	126.813350
4	광주송정도서관	광주광역시	광산구	공공도서관	광주광역시교육청	062-940-8900	http://lib.gen.go.kr/songjung/	35.145316	126.799604
5	첨단도서관	광주광역시	광산구	공공도서관	광주광역시 광산구청	062-960-8296	http://lib.gwangsan.go.kr	35.217233	126.830102
6	장덕도서관	광주광역시	광산구	공공도서관	광주광역시 광산구청	062-960-3989	http://gwangsan.lib.go.kr	35.187519	126.813380
7	이야기꽃도서관	광주광역시	광산구	공공도서관	광주광역시 광산구청	062-960-6812	http://gwangsan.lib.go.kr	35.148038	126.780524
8	첨단도서관	광주광역시	광산구	공공도서관	광주광역시 광산구청	062-960-8297	http://gwangsan.lib.go.kr	35.216792	126.832239
9	신가도서관	광주광역시	광산구	공공도서관	광주광역시 광산구청	062-960-8294	http://gwangsan.lib.go.kr	35.183484	126.824308
10	운남어린이도서관	광주광역시	광산구	공공도서관	광주광역시 광산구청	062-960-8302	http://lib.gwangsan.go.kr/	35.173025	126.819786

```
In [25]: gmaps_key = "AIzaSyDjqdtQBjvQVHYI3bWENsC7qU7v77nL144"
gmaps = googlemaps.Client(key=gmaps_key)

mapping = folium.Map(location=[test['위도'].mean(), test['경도'].mean()], zoom_start=11)

for n in test.index:
    folium.Marker([test['위도'][n], test['경도'][n]]).add_to(mapping)

mapping
```



```
In [22]: location = pd.read_excel('../data/lotation.xls', header = 0)
location.fillna(method='pad', inplace=True)

location = location[location['시도명'] == '광주광역시']
location = location[location['시군구명'] == '동구']
location.reset_index(drop=True, inplace=True)

location[['도서관명', '시도명', '시군구명', '도서관유형', '운영기관명', '도서관전화번호', '홈페이지주소', '위도', '경도']]
```

	도서관명	시도명	시군구명	도서관유형	운영기관명	도서관전화번호	홈페이지주소	위도	경도
0	극립아시아문화전당 라이브러리파크	광주광역시	동구	전문도서관	아시아문화원	1899-5566	http://library.acc.go.kr	35.147218	126.920666
1	산수도서관	광주광역시	동구	중증도서관	광주광역시립도서관	062-613-7831	http://www.citylib.gwangju.kr/main/main.php	35.154691	126.928434
2	계림꿈나루도서관	광주광역시	동구	중증도서관	광주광역시 동구청	062-608-3920	http://lib.donggu.kr	35.158532	126.919653
3	중앙도서관	광주광역시	동구	중증도서관	광주광역시교육청	062-607-1300	http://lib.gen.go.kr/jungang/	35.148503	126.925464
4	식물그림책작은도서관	광주광역시	동구	작은도서관	기살림빛고을협동조합	062-224-8863	http://lib.namgu.gwangju.kr/	35.123369	126.907744
5	그림이 있는 작은도서관	광주광역시	동구	작은도서관	개인	070-8723-8834	http://lib.namgu.gwangju.kr/	35.123369	126.907744
6	GIC도서관	광주광역시	동구	작은도서관	(사)광주국제교류센터	062-226-2733	http://lib.namgu.gwangju.kr/	35.123369	126.907744
7	space秀	광주광역시	동구	작은도서관	개인	062-223-4127	http://lib.namgu.gwangju.kr/	35.123369	126.907744
8	지산1동 작은도서관	광주광역시	동구	작은도서관	지산1동 복지공동체	062-608-3682	http://lib.namgu.gwangju.kr/	35.123369	126.907744
9	작은도서관 오월의술	광주광역시	동구	작은도서관	(사)합수운한봉기념사업회	062-514-0528	http://lib.namgu.gwangju.kr/	35.123369	126.907744
10	무등재 작은도서관	광주광역시	동구	작은도서관	개인	062-233-2489	http://lib.namgu.gwangju.kr/	35.123369	126.907744
11	두암타운작은도서관	광주광역시	동구	작은도서관	두암타운입주자대표회의	062-266-3784	http://lib.namgu.gwangju.kr/	35.123369	126.907744
12	소망작은도서관	광주광역시	동구	작은도서관	개인	062-430-2558	http://lib.namgu.gwangju.kr/	35.123369	126.907744
13	이매진도서관	광주광역시	동구	작은도서관	개인	000-000-0000	http://lib.namgu.gwangju.kr/	35.123369	126.907744
14	회랑작은도서관	광주광역시	동구	작은도서관	횡성한교회	000-000-0000	http://lib.namgu.gwangju.kr/	35.123369	126.907744
15	계림꿈나루도서관	광주광역시	동구	중증도서관	광주광역시 동구청	062-234-2293	http://blog.naver.com/dreamprime	35.158532	126.919653
16	지원2동작은도서관	광주광역시	동구	작은도서관	광주광역시 동구청	062-608-3922	http://blog.naver.com/dreamprime	35.120452	126.932873
17	지산2동작은도서관	광주광역시	동구	작은도서관	광주광역시 동구청	062-225-2123	http://blog.naver.com/dreamprime	35.148426	126.938491
18	학운동작은도서관	광주광역시	동구	작은도서관	광주광역시 동구청	062-608-3921	http://blog.naver.com/dreamprime	35.132657	126.937004
19	대인작은도서관	광주광역시	동구	작은도서관	대인교회	062-234-8291	http://blog.naver.com/dreamprime	35.132657	126.937004
20	계림두산위브작은도서관	광주광역시	동구	작은도서관	계림두산위브아파트입주자대표회의	062-227-0338	http://blog.naver.com/dreamprime	35.132657	126.937004
21	행복한작은도서관	광주광역시	동구	작은도서관	행복한지역아동센터	062-223-3164	http://blog.naver.com/dreamprime	35.132657	126.937004
22	무등산IPARK아파트작은도서관	광주광역시	동구	작은도서관	무등산IPARK입주자대표회의	062-229-0116	http://blog.naver.com/dreamprime	35.132657	126.937004
23	알암인권작은도서관	광주광역시	동구	작은도서관	(사)알암영노근선상기념사업회	062-228-3563	http://blog.naver.com/dreamprime	35.132657	126.937004
24	푸른길책활기작은도서관	광주광역시	동구	작은도서관	(사)동명창조문화협복공동체마을	062-233-2662	http://blog.naver.com/dreamprime	35.132657	126.937004
25	대안도서관 틈을	광주광역시	동구	작은도서관	광주전남 인드라망상명공동체	062-268-4660	http://blog.naver.com/dreamprime	35.132657	126.937004
26	산수종양작은도서관	광주광역시	동구	작은도서관	산수종양교회	062-251-6438	http://blog.naver.com/dreamprime	35.132657	126.937004
27	풀꽃동산작은도서관	광주광역시	동구	작은도서관	지산교회	062-222-0736	http://blog.naver.com/dreamprime	35.132657	126.937004
28	북카페 영어작은도서관	광주광역시	동구	작은도서관	광남교회	062-234-9182	http://blog.naver.com/dreamprime	35.132657	126.937004
29	백암문고	광주광역시	동구	작은도서관	백암지역아동센터	062-227-9421	http://blog.naver.com/dreamprime	35.132657	126.937004
30	파이디온작은도서관	광주광역시	동구	작은도서관	파이디온지역아동센터	062-265-5616	http://blog.naver.com/dreamprime	35.132657	126.937004
31	은새암	광주광역시	동구	작은도서관	(재)광주광역시 광주여성재단	062-670-0539	http://blog.naver.com/dreamprime	35.132657	126.937004
32	작은도서관 숲	광주광역시	동구	작은도서관	개인	062-401-1708	http://blog.naver.com/dreamprime	35.132657	126.937004
33	메트로열린도서관	광주광역시	동구	작은도서관	광주문화사랑봉사단	062-604-8434	http://blog.naver.com/dreamprime	35.132657	126.937004
34	율달샘작은도서관	광주광역시	동구	작은도서관	개인	062-233-0433	http://blog.naver.com/dreamprime	35.132657	126.937004
35	학2마을작은도서관	광주광역시	동구	작은도서관	학2마을아파트입주자대표회의	062-223-1168	http://blog.naver.com/dreamprime	35.132657	126.937004
36	율남호반동서지유도서관	광주광역시	동구	작은도서관	개인	062-227-0660	http://blog.naver.com/dreamprime	35.132657	126.937004
37	슬레이시오 작은도서관 마인	광주광역시	동구	작은도서관	개인	062-233-1013	http://blog.naver.com/dreamprime	35.132657	126.937004
38	예향작은도서관	광주광역시	동구	작은도서관	개인	062-232-7429	http://blog.naver.com/dreamprime	35.132657	126.937004
39	김넷작은도서관	광주광역시	동구	작은도서관	개인	062-229-3355	http://blog.naver.com/dreamprime	35.132657	126.937004

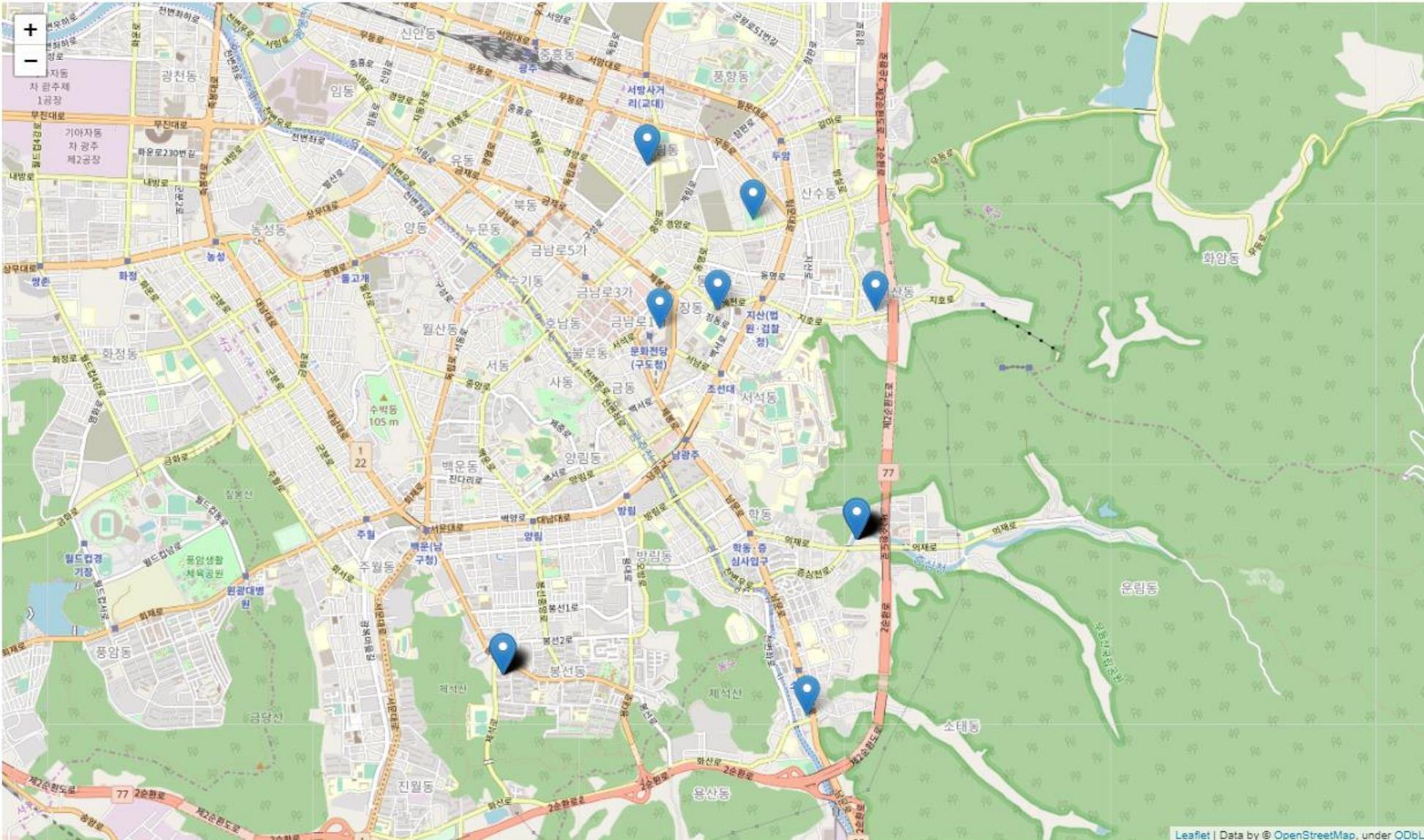
```
In [23]: gmaps_key = "AIzaSyDjqdtQBjvQVHYI3bWENsC7qU7v77nL144"
gmaps = googlemaps.Client(key=gmaps_key)

mapping = folium.Map(location=[location['위도'].mean(), location['경도'].mean()], zoom_start=11)

for n in location.index:
    folium.Marker([location['위도'][n], location['경도'][n]]).add_to(mapping)

mapping
```

Out [23]:



고찰/소감

지금까지 우리나라 독서율의 실태를 조사하면서 알게된점은 크게 4가지입니다.

- 1) 우리의 여가활동 리스트중에서 독서라는 비중이 계속해서 낮아지고 있다는 점.
- 2) 원인은 시간에 치여 책을 읽을 시간이 없고, 책 이외에도 우수한 매체들이 많기 때문이라는 것.
- 3) 그로인해 계속해서 독서율이 낮아지게 될 것이라는 것.
- 4) 독서율을 증가시킬 수 있는 좋은 방법은 도서관 이용률을 증가시키면 된다는 것.

계속해서 낮아지고 있는 독서율은 우리나라에만 해당되는 문제가 아니라 세계적인 이슈라고 합니다. 하지만 아무리 노력해도 종이책의 독서율은 정보화 시대의 발전에 따라 계속해서 낮아질 수밖에 없습니다. 전자책을 보면 되는 것이 아닌가라는 의문을 갖는 사람도 있겠지만, 아직 종이책에 비해 전자책의 출간 수는 턱없이 부족하다고 합니다. 특히 우리나라는 외국 시장에 비해 상당히 뒤쳐져 있으며 애초에 책 자체를 잘 읽지 않아서 특정 분야의 도서만 출판되고 있는 실정이라고 합니다. 물론 전자책은 종이책에 비해 많은 이점이 있지만 그럼에도 불구하고 아직까지는 전자책이 완벽하게 종이책의 대체제가 될 수 없을 것 같습니다. 독서를 남 주기 위해서 하는것은 아니지만 정부차원에서 문제가 되는걸 인지하고 있다면 해당 문제에 대해 위기감과 책임감을 갖고 여러 방향으로 독서율을 늘리기 위한 노력을 보여줬으면 좋겠습니다. 예를 하나 들자면, 현재 독서율은 노인층으로 갈수록 줄어들고 있는 실정입니다. 당장 저희 부모님은 노안으로 스마트폰의 기본 폰트조차 잘 보이지 않아서 가장 큰 폰트 크기로 바꿔 사용중입니다. 이런 분들이 책을 읽기란 곤란할 것입니다. 어른들을 대상으로 안경점과 제휴해서 돋보기를 나눠주는 등 다양한 지원을 해주고, 독서 이벤트를 진행하는 방안 등이 있었으면 좋겠습니다.

참고

- 독서의 중요성

- 1) <https://news.samsung.com/kr/책-안-읽는-시대-그럼에도-불구하고-독서가-필요>
- 2) <https://blog.naver.com/eggshamegg/220783800562>

- 여가활동별 참여비율 (2016~2018)

http://kosis.kr/statHtml/statHtml.do?orgId=113&tblId=DT_113_STBL_1028375&conn_path=I3

- 여가활동별 참여비율 (2006~2008)

http://kosis.kr/statHtml/statHtml.do?orgId=113&tblId=DT_113_STBL_1014277&conn_path=I3

- 독서량(2009~2017)

http://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=DT_1SSCL020R&conn_path=I3

- 독서 장애 요인 변화 추이

http://kosis.kr/statHtml/statHtml.do?orgId=113&tblId=DT_113_STBL_1024028&conn_path=I3



참고

- 성인의 인근지역 공공도서관 이용률 및 이용 빈도

http://kosis.kr/statHtml/statHtml.do?orgId=113&tblId=DT_113_STBL_1024369&conn_path=l3

- 성인이 공공도서관을 이용하지 않는 이유

http://kosis.kr/statHtml/statHtml.do?orgId=113&tblId=DT_113_STBL_1015728&conn_path=l3

- 종이책+전자책 독서율

http://kosis.kr/statHtml/statHtml.do?orgId=113&tblId=DT_113_STBL_1025546&conn_path=l3

- 종이책 독서율

http://kosis.kr/statHtml/statHtml.do?orgId=113&tblId=DT_113_STBL_1025544&conn_path=l3

- 공공 데이터 포털 공공 도서관 위치

<https://www.data.go.kr/dataset/15013109/standard.do>



이번 프로젝트에서는 제가 직접 개인 프로젝트의 주제를 정하고 기능을 어떻게 구현할지 처음에 설계하면서 여러 가지 기능들을 구현하려고 생각했었는데, 프로그램을 구현하면서 상상했던것만큼 구현할 수 없었습니다. 책에서 나온 기능을 최대한 사용하면서 프로젝트를 진행하고 싶었는데, 시간적으로나 능력적으로나 부족한 부분이 많아서 몇가지는 구현을 포기하고 간단한 기능으로 대체한 것이 아쉬웠습니다.

3장의 웹페이지에 대한 정보를 읽어와서 위치를 지도에 표기하는 것을 해보고 싶었는데 놀랍게도 공공 데이터 센터에서 전국의 도서관별 위도와 경도를 제시한 데이터가 있어서 간단하게 그 위도와 경도를 사용해서 지도에 위치를 표시하는 것으로 대신했던 점과, 7장의 시계열 데이터를 꼭 사용해보고 싶었는데 데이터가 너무 부족해서 진행할 수 없었던 애로사항이 있었습니다.

이번학기 프로젝트는 제가 PYTHON 언어에 대한 이해도가 부족해서 프로젝트 진행에 걱정이 많았었는데, 막상 만들어서 결과물을 보니 개인적으로는 만족스러운 결과가 나온 것 같습니다.

기말고사 프로젝트 Fin.

교수님 한학기동안 고생하셨습니다!

